

医学系対訳コーパスを使用した英語学習 Web アプリケーション —学習者のスキル向上を指向した機能拡張—

Web Application to Help Learn English Using Medical Bilingual Corpus - Extensions to Improve Learners' Skills -

安田 陸^{*1}, 宮崎 佳典^{*2}, 浅野 元子^{*3}, 藤枝 美穂^{*3}

Riku Yasuda^{*1}, Yoshinori Miyazaki^{*2}, Motoko Asano^{*3}, Miho Fujieda^{*3}

^{*1} 静岡大学情報学部情報科学科

^{*1}Faculty of Informatics, Shizuoka University

^{*2} 静岡大学大学院情報学領域

^{*2}College of Informatics, Shizuoka University

^{*3} 大阪医科薬科大学医学部

^{*3}Faculty of Medicine, Osaka Medical and Pharmaceutical University

Email: yasuda.riku.20@shizuoka.ac.jp

あらまし: 医学教育分野においては、英語論文の読解や執筆には分野依存の表現に深い理解が必要である。これに対し、我々は医学系対訳コーパスを活用した英語学習支援 Web アプリケーション MEESUS (Medical English Education Support System) の開発を進めている。本発表では、英文読み上げ機能の追加による学習者のリスニングスキル向上、学習者が作成した英文に対して類似文を提示するライティングスキル支援を目指す。パイロット実験で得られた好意的な評価についても触れる。

キーワード: 医学英語、対訳コーパス、Web アプリケーション、英文読解・作文支援

1. はじめに

医学教育の分野では、臨床カンファレンスなどで医学の最新情報を得る、医療の最新情報を得るために英語論文を読むなど、英語を効果的に活用する能力が特に重要である。また、医師・医学専門教員を対象とした英語能力のニーズ調査から、学生に求めるスキルとして論文読解のニーズが高いことが分かっている⁽¹⁾。これは ESP (English for Specific Purposes, 特定の目的のための英語) の中でも特に医学分野に特化した EMP (English for Medical Purposes) の領域を意味する。これらの背景から本研究室では提供された医学英語論文 (タイトルならびに概要部) の対訳付コーパスに対し、医学英語学習者を対象に EMP 学習支援を目的とした MEESUS⁽²⁾ (以下、本システム) を開発した。現段階の本システムは英文読解・英作文支援を視野に入れているが、幅広い EMP 学習支援を実現するため、本研究ではさらにリスニング・ライティング支援に向け拡張することを目的とする。

関連研究には JECPRESE⁽³⁾ や AWSuM⁽⁴⁾ などが挙げられる。⁽³⁾ は KWIC (KeyWord In Context) 表示機能を備えており、口頭発表をテキスト化したデータを対象に話者の目的に応じて英語と日本語の文を検索する。⁽⁴⁾ はムーブ (表現意図) を活用した英文作成支援システムであり、論文のセクション単位でムーブごと高頻度の語連鎖を提示する。コンコーダ機能も提供されているが、日本語対訳表示はサポートされていない。

2. 医学英語学習システム

2.1 医学英語論文コーパス

以下の図 1 に、医学英語論文のタイトルおよび概要部に対訳が付いたコーパスのデータ構造を示す。なお、このタイトル・概要部の 1 つを英文書と規定する (ここには英文書 3,920 冊、英文数 49,884 が含まれており、各英文は日本語対訳を有する)。英文書はタイトル、背景、手法、結果、結論 (研究情報) の 6 セクション構成である。図 1 では ID=1695 の英文が英文書 2 の 14 番目の文で、結論部で使用されていることが示されている。

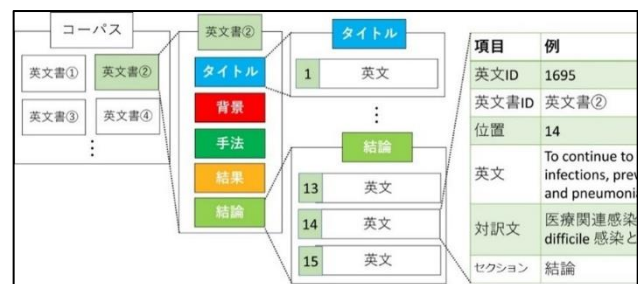


図 1 データ構造, データ例

2.2 医学英語論文コーパス

本システムの概要を図 2 に示す。検索方法として単一検索、複合検索、検索窓それぞれに日本語と英語を組み合わせたクロス検索がある。検索オプションには文章構成要素のセクション指定、KWIC 表示用の語数指定、活用表現等の詳細設定機能が備わっている。検索エンジンは Elasticsearch を使用し、ユーザ入力から検索クエリを自動生成する。検索結果はヒットした英文単位で KWIC 表示される。その後、英文と対訳文全体を表示する機能、ヒットした英文が

属する英文書全体を表示する機能などを使用できる。各種機能の操作時には学習者の操作ログを取得し Learning Analytics に関連付けることを主眼におく。

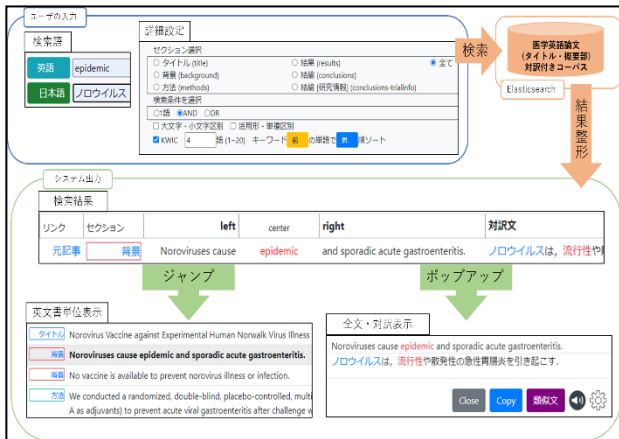


図2 システムインタフェース

3. 評価実験・分析

国内医科薬科大学の学生 112 名に対して授業に実験を実施し、使用履歴を分析後、システムの改善点・拡張機能を検討した。アンケート内「システムを使用するとアカデミックな文章を読むことの難しさが減じると感じられますか」の回答結果として概算で（高評価, 低評価, 特になし）= (76%, 5%, 19%) の比率が得られた。同様に本システムの使用感に対する回答の低評価より具体例（抜粋）を表1に示す。

表1 アンケート回答 低評価内容

Web 辞書, 翻訳のほうが発音も聞いて使いやすい
自分の希望している形の文章が上の方に表示されない
Google 翻訳でいいかなと思った
機能的には機械翻訳のほうが便利

以上より、論文抄録と対訳を用いたリーディングスキルを支援する部分が評価された一方で、他の辞書などと比して機能面で劣る可能性も指摘された。これに対し、従前機能に加えリスニングおよびライティングスキルを効果的に学ぶ機能の実装を行う。

4. 追加機能

4.1 英文読み上げ機能

英文読み上げ機能は Web Speech API⁽⁵⁾ を活用して検索結果の英文を読み上げ、リスニングスキル向上を目指す。しかし一般的な発音付き辞書には掲載されない医学専門用語の発音を保証できない点がある。代表的な3つの辞書に現れない単語を抜き出した結果、コーパス内 1,616 語（全体の 6.6%）が該当した。そこで該当する医学用語に注釈をつけることで、不明な発音に対する注意喚起を与える機能を実装した。

4.2 類似文検索機能

類似文検索機能は EWSS⁽⁶⁾ を用いて検索結果やユーザによる入力英文の類似文を提示し、ライティン

グスキル向上を目指す。入力文と英文書のコーパス間で類似度が計算され降順に上位 50 件が提示

1	The Spingomonas glycolipids stimulated the release
2	PSTs cause human fatalities by blocking sodium cond
3	Various types of constraints cause this .
4	Question What causes posttraumatic stress disorder ?
5	Neither of these species causes plant disease .
6	John 's overeating causes his obesity .
7	Thus , incorrect presuppositions cause two problems

図3 類似文検索結果 データセットは元々

情報学・科学系のものであるため、MEESUS で使われている文献を適応させ、医学専門用語に優先的に重みづけを行うことで EMP 用に昇華させる。

4.3 類似文検索機能に関するパイロット評価実験

某大学の情報系学部生・大学院生計 10 名に試用してもらい操作性や有効性に関するアンケートを実施した。結果、ライティングスキル支援ツールとして好意的な感想「英語文の書き換えにとっても便利」「(自動生成ではなく)実在する英文を提示するのが良い」などが寄せられ、90%が本機能によりアカデミック文書を書く難易度が減じると答えた（同じく実施した英文読み上げ機能については紙面の都合上割愛）。

5. まとめ

本研究では、現行のシステムで実験を行い、そこで得られたアンケート結果を分析し、2つの追加機能の実装を行った。英文読み上げ機能では、正しい発音を保証できない専門用語に注釈をつけることで、質の高いリスニングスキル支援の提供を行う。類似文検索機能では、医学用語に重みづけを行い、より医学に特化した類似文を提示することで、専門性の高いライティングスキル支援の提供を行う。パイロット実験の結果を発表時に報告する予定である。

参考文献

- (1) Naruenatwatana, N., Vijchulata, B., A study of the needs of medical students in the use of academic English perceived by three groups: Medical students, teachers of English and subject teachers, Studies in Language and Language Teaching, 10, pp. 1–23 (2001).
- (2) M. Asano, M. Nakano, Y. Miyazaki, M. Fujieda, Introducing a Bilingual Corpus Database System of Medical Abstracts for Exploring Academic Connotations of Words: A Case Study of First-year Medical Students, J. Med. Eng. Educ., 21 (1), pp. 18-26, (2022).
- (3) Kunioishi, et al., An online support site for preparation of oral presentations in science and engineering, European Journal of Engineering Education, 37:6, 600–608 (2012).
- (4) 水本篤, 浜谷佐和子, 今尾康裕, ムーブと語連鎖を融合させたアプローチによる応用言語学論文の分析—英語学術論文執筆支援ツール開発に向けて—, 英語コーパス研究 23, pp. 21-32 (2017) .
- (5) Web Speech API, https://developer.mozilla.org/en-US/docs/Web/API/Web_Speech_API (参照 2023-12-22).
- (6) Y. Miyazaki, S. Tanaka, Y. Koyama, Development of a Corpus-Based Web Application to Support Writing Technical Documents in English, E-Learn 2014, Vol. 2014, No.1, pp.1371-1380 (2014).